

Statistical tests for group comparison of manifold-valued data

Anne Collard, Christophe Phillips and Rodolphe Sepulchre

Abstract—Motivated by population studies of Diffusion Tensor Imaging, the paper investigates the use of mean-based and dispersion-based permutation tests to define and compute the significance of a statistical test for data taking values on nonlinear manifolds. The paper proposes statistical tests that are computationally tractable and geometrically sound for Diffusion Tensor Imaging.

I. INTRODUCTION

Statistical analysis of scalar-valued images is a well established central component of contemporary science. But the evolution of sensor technology and data storage increasingly produces images that are multimodal and nonlinear in nature. This has motivated significant work in the recent years to extend signal processing techniques from scalar-valued data to manifold-valued data, see e.g. [1].

The present paper is specifically motivated by the increasing role of diffusion tensor imaging in neuroscience. In their simplest form, DT images provide a 3×3 positive definite diffusion tensor for each voxel [2]. Voxel-based statistical analysis of a population therefore involves statistical tests among positive definite tensors rather than scalars. The default remedy is to convert the tensor information into a scalar information (usually fractional anisotropy, see below), but both the intensity information (that is, the three positive eigenvalues of the tensor) and the orientation information (the orientation of the three principal axes) potentially contain valuable statistical information, calling for new methodological developments.

The challenge is methodological as well as computational because clinical studies usually involve large populations and many voxels, that is, large-scale statistical analysis.

With this motivation in mind, the present paper investigates the methodological and computational value of two standard non-parametric permutation tests: a mean-based permutation test and a dispersion-based permutation test. We discuss how these tests can be extended from scalar-valued data to manifold-valued data and specialize the discussion in the particular case of Diffusion Tensors, that is, 3×3 positive definite matrices.

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its author(s).

A. Collard is with the Departement of Electrical Engineering and Computer Science, University of Liège, 4000 Liège, Belgium; anne.collard@ulg.ac.be

R. Sepulchre is with the Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom and with the Departement of Electrical Engineering and Computer Science, University of Liège, 4000 Liège, Belgium; r.sepulchre@eng.cam.ac.uk

C. Phillips is with the Cyclotron Research Centre, University of Liège, 4000 Liège, Belgium; c.phillips@ulg.ac.be

We argue that a dispersion-based permutation test is a computationally tractable approach to clinical DTI statistical analysis and stress the value of defining properly motivated geometric quantities for the underlying similarity measures.

A. State of the art for statistical analysis of DTI

There is a vast recent literature on DTI statistical analysis. Among these papers, we can distinguish the ones using univariate tests to compare DT images and the ones trying to exploit further information. For example, [3], [4], and [5] focused on the fractional anisotropy or the mean diffusivity of tensors, without paying any attention to the orientation of these tensors. These scalars are indeed invariants linked to the shape of tensors, but can not detect any difference in orientation between tensors. In [6], the evolution of the fractional anisotropy along a fiber tract is studied. Some other papers have tried to use the whole information contained in the tensor, namely through the use of the Log-Euclidean metric (*i.e.* they use the logarithms of tensors instead of the tensors themselves). This is the case in [7], where an Hotelling's T^2 test is developed for the Log-Euclidean metric, a framework similar to the ones in [8] and in [9]. Some other papers presented a rigorous conceptual framework based on the Riemannian manifold $S^+(3)$, as in [10], [11] and [12]. However, those papers have not addressed the statistical significance of a test for group comparison. This is also the case of [13], which uses another parametrization of tensors. Statistical tests were proposed in [14], through a decoupled analysis of the eigenvalues and the eigenvectors of the tensors. The tests are based on distributional assumptions, which is a potential limitation for diffusion tensors. The closest published work to the present paper is [15], when the authors propose a multivariate dispersion-based permutation test, see Section III for more details.

The paper is organized as follows: after a brief review of the state of the art for the statistical analysis of Diffusion Tensor, Section II will focus on statistical tests for group comparison, beginning with mean-based permutation tests. Computation of means on manifolds will also be discussed, before the introduction of dispersion-based tests. The case of multivariate tests will be addressed, and some appropriate similarity measures for DTI will be introduced. Section III deals with the methods that we have used for our tests, while Section IV shows several of our results.

II. STATISTICAL TESTS FOR GROUP COMPARISON

Statistical analyses of scalar images are often performed through the use of parametric methods [16], such as the Student t^2 test for comparison of Gaussian variables. However, the distribution of multivariate data is rarely known and, if known, is often not Gaussian. This explains why many authors have made the choice of non-parametric methods to study multivariate images. Among these non parametric methods, permutation tests are often used because of their relative simplicity. Permutation methods provide statistical significance testing of difference between groups without having to assume a distribution of the data [17]. These methods have the ability to directly estimate the null distribution of the statistics describing the difference. Moreover, these methods are easily applicable to any statistical test, which is interesting to compare results obtained with different parametrizations of the data.

A. Mean-based permutation tests

Permutation tests are based on a simple idea. For the sake of illustration, consider the statistical significance of a variable x to distinguish among two populations C and D. Suppose that the mean $\mathbb{E}(x)$ differs by a quantity Δ_0 between group C and group D. Permutation tests enable to quantify, without assumptions about the distribution of the variable, if this difference is significant or not. Indeed, if the difference is not significant, it should not be altered by random permutations between C and D. Therefore, given the null hypothesis that the labelings are arbitrary, the significance of x can be assessed by comparison with the distribution of values under all possible permutations. This is illustrated by the histogram in Figure 1. If the observed difference Δ_0 is in the tail of the distribution, it means that very few permutations of the data attain the same difference. The p -value of the test is given by the ratio between the number of times that a permuted statistics is higher than the observed value and the number of performed permutations. The test is statistically significant at a level α if the p -value is smaller than α . This means that this value has less than $100\alpha\%$ of chance to have be found randomly.

The generalization of a permutation test to data that take values on a manifold is conceptually straightforward because it only requires a proper notion of mean. On a Riemannian manifold \mathcal{M} , the Karcher mean of a set of points $\{x_1, x_2, \dots, x_N\}$ is given by the Fréchet formula

$$\mu = \arg \min_{x \in \mathcal{M}} \frac{1}{2N} \sum_{i=1}^N d^2(x, x_i), \quad (1)$$

(with d the distance on the manifold), which reduces to the classical arithmetic mean when using the Euclidean distance.

B. Means on manifolds and means for DTI

Computing Riemannian means (or medians) on specific manifolds has been the object of significant research in the recent years, see e.g. [18] for means on the Grassmann

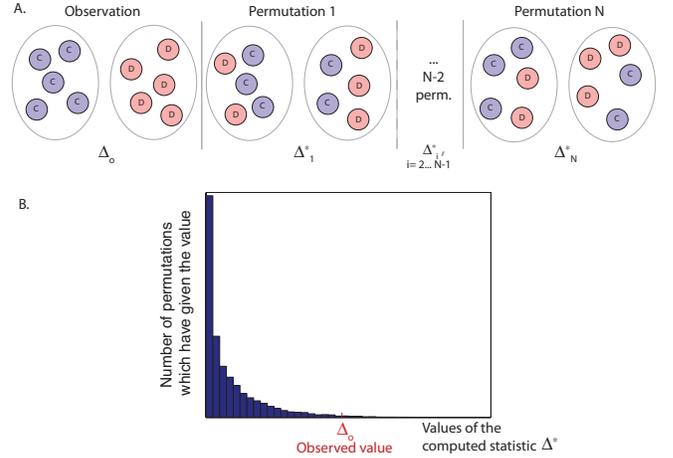


Fig. 1. General procedure of a permutation test. Given the observed value of the statistic, values corresponding to N permutations of the labels are computed. This gives an approximation of the (unknown) distribution of the statistic. The comparison between the observed value and this distribution enables to compute the p -value of the test.

manifold, [19], [11] for means and medians on the space of positive definite matrices, and [20] for a mean for fixed-rank positive semidefinite matrices.

It should be emphasized that a mean-based permutation test may represent a formidable computational task. For instance, computing the Riemannian mean of several positive definite matrices is typically achieved by an iterative algorithm [19]. For a population of size N classified in G groups of size n_g , this computation must be repeated $M = N! / \prod_{g=1}^G n_g!$ which is prohibitive for large populations, even if the full distribution is not computed. In the context of DTI, the statistics are typically computed for a large number of voxels, which adds to the computational burden.

One computational remedy for positive definite tensors is to compute the arithmetic mean of tensors (that is, to work with the Euclidean metric) or, better, to compute a matrix geometric mean according to the formula

$$\mu_{LE} = \exp\left(\sum_{i=1}^N \log x_i\right). \quad (2)$$

which corresponds to the Riemannian mean for the Log-Euclidean metric [21].

In a recent paper, the authors have introduced another notion of mean for positive tensors, that provides a better decoupling between orientation and anisotropy [22]. Using the spectral decomposition of each tensor, the mean tensor is defined as

$$\mu_{SQ} = R_\mu \Lambda_\mu R_\mu^T, \quad (3)$$

where Λ_μ is the geometric mean of the eigenvalues of the tensors, i.e. $\Lambda_\mu = \exp(\sum_{i=1}^N \log \Lambda_i)$, where $\Lambda_i = (\lambda_{1,i}, \lambda_{2,i}, \lambda_{3,i})$, with the ordered eigenvalues $\lambda_{1,i} > \lambda_{2,i} > \lambda_{3,i}$. The mean orientation R_μ is computed through the chordal mean of quaternions.

Even with such computational simplifications, the computational burden of a mean-based approach remains prohibitive for a real application of group studies using Diffusion Tensor Images because the matrix log operation will scale in a factorial way with the population size.

C. Dispersion-based permutation tests

The Multiresponse Permutation Procedure (MRPP) proposed in [23] is not based on repeated computations of means but only requires to build a similarity matrix between all data points. Given a symmetric similarity measure $s(i, j)$ between data points x_i and x_j , the similarity matrix is defined as the symmetric matrix S with element $S_{ij} = s(i, j)$.

The statistical test proposed in [23] is based on a measure of dispersion of the data points within each group rather than on a measure of mean. The dispersion δ_g of a group g of n_g elements is defined as

$$\delta_g = \frac{2(n_g - 2)!}{n_g!} \sum_{\substack{i < j \\ i, j \in g}} s(i, j) \quad (4)$$

where the sum is computed over all data points of the group.

The overall dispersion δ of the variable x in the population is defined as a weighted sum of the dispersions in each group:

$$\delta = \sum_{g=1}^G C_g \delta_g$$

where $C_g > 0, g = 1 \dots G$ are the weights of each group (their sum must be 1).

The rest of the procedure follows the permutation test described in Section II. A., with the mean-difference Δ replaced by the dispersion δ . The observed dispersion is judged statistically significant only if it occurs in the lower tail of the histogram among all possible permutations of the population.

The distance-based permutation test has the same advantages than the mean-based permutation test: it does not require any assumption about the statistical distribution and only requires a similarity measure between data points, for instance a distance on Riemannian manifolds. But a significant computational advantage is that the similarity matrix must be computed only once, requiring $O(N^2)$ computations of pairwise similarity for a total population of size N .

D. Multivariate testing

Our description of permutation tests has so far assumed an univariate statistical testing but is easily extended to multivariate testing. A widely used method for the statistical comparison of multivariate data consists in the computation of 'marginal' or 'partial' tests (one for each of the n_v considered variable) and to combine the p -values obtained for each partial test by a combining function, which can be of different forms [24]. This method involves two stages of computation. First, the marginal p -values (denoted $\xi_i, i = 1, \dots, n_v$) are computed through a permutation test (here, we will use a MRPP test for univariate data). Then, the combining function is used to compute the combined observed value,

$T_o = \mathcal{C}(\xi_1, \dots, \xi_{n_v})$. The distribution of this T is computed through the combination of the p -values computed for each permutation of the first step, *i.e.* $T_k^* = \mathcal{C}(\xi_{1,k}^*, \dots, \xi_{n_v,k}^*)$. The combined p -value for the test is then estimated via the ratio between the number of occurrences where T_k^* was larger than T_o and the total number of performed permutations.

E. Similarity measures for DTI

Every distance on the manifold of positive tensors qualifies for a similarity measure. In the present work, we will compare two similarity measures to the naive Euclidean distance between matrices: the Log-Euclidean distance [21] and the spectral-quaternion similarity measure recently introduced in [22]. Both similarity measures only involve a spectral decomposition of a 3×3 matrix for all data points as the main computation of the similarity measure.

Regarding multivariate testing, we propose to compare an "Euclidean" statistical test based on the six independent quantities of a 3×3 positive definite tensor (as proposed in [15]) and a "geometric" statistical test based on the six geometric quantities that define scaling and orientation of the tensor: the three eigenvalues and the three first components of the quaternion. In the latter case, we use a geometric similarity measure for the (positive) eigenvalues $s(\lambda, \mu) = \sqrt{\log^2 \left(\frac{\lambda}{\mu} \right)}$ and an Euclidean (chordal) measure for the quaternions. A further alternative would be a 'Log-Euclidean' statistical test based on the six elements of the logarithm of the tensor, as investigated in [15]. We do not include this comparison here since the work of [15] suggests that it does not offer significant advantages compared to the Euclidean test.

For comparison purposes, we will also compute an univariate test using the fractional anisotropy of tensors [4].

III. METHODS

We will compare the power of the proposed statistical tests on the following synthetic data sets.

We generate two groups of tensors starting from a reference tensor and a transformed tensor. The parameter γ quantifies the amount of difference in a specific parameter between the two groups. Denoting by $\lambda_i, i = \{1, 2, 3\}$, the eigenvalues of the reference tensor and by λ'_j the ones of the deformed tensor, the following five different geometric transformations of the reference diffusion tensor are studied [25]:

- *Decrease of longitudinal diffusion (DL)*: λ'_1 will be given by $\lambda'_1 = \lambda_1 - \gamma(\lambda_1 - \lambda_2)$, $\gamma \in [0, 1]$. The two others eigenvalues will be left unchanged.
- *Increase of radial diffusion (IR)*: λ_2 and λ_3 will be replaced by $\lambda'_j = \lambda_j + \gamma \frac{\lambda_1 - \lambda_j}{2}$, $j = \{2, 3\}$, $\gamma \in [0, 1]$, while λ_1 will be left unchanged.
- *Increase of mean diffusion (IM)*: all the eigenvalues of the deformed tensor will be given by $\lambda'_j = (1 + \gamma)\lambda_j$, $j = \{1, 2, 3\}$, $\gamma \in [0, 1]$.

TABLE I
PARAMETERS OF THE STATISTICAL TESTS

Parameter	Value
Level of significance α	0.05
Number of samples by groups n_i	10
Number of permutations N_p	20 000
Number of tests by situation N_t	500

- *Change of diffusion orientation (CO)*: the angle θ between principal directions of the reference tensors will change following $\theta' = \theta + \gamma \frac{\pi}{2}$, $\gamma \in [0, 1]$.
- *Change in both eigenvalues and orientation*: in this case, we will combine a difference in orientation (CO) with one of the first three differences. Both modifications will evolve following the same γ .

Starting from the reference tensor and its reference deformation, we generate a population of $N = 20$ tensors by sampling from a Wishart distribution 10 tensors around the reference tensor and 10 tensors around the deformed tensor.

The statistical comparison will be tested in a situation of high anisotropy, where the eigenvalues of the reference are $\Lambda = (5, 1, 0.5)$, a situation of small anisotropy where they are given by $\Lambda = (3, 1, 1)$, and a situation of near isotropy, $\Lambda = (1.3, 1, 1)$. The performance of the different tests in many situations will be assessed through the computation of the *power* of these tests. This quantity is computed by performing the same test a large amount of times and by counting the occurrences of significant results. The power is the ratio between this number of occurrences and the total amount of performed tests. If the test is not efficient, the power will be close to the significance level α . A value of the power equal to one means that the test is particularly efficient for the analyzed situation. The different used values (number of tests, level of significance, ...) are summarized in Table I. It should be noted that the level of noise is relatively high for these tests. The tests could be done with a larger number of degree of freedom in the Wishart distribution (which corresponds to a lower level of noise).

IV. RESULTS

In the following, some interesting results found during our tests will be shown. We will first study how dispersion-based tests vary with the used similarity measure and then observe the difference in interpretation of the results which can be done from multivariate tests. Due to space constraints, other results will not be shown.

A. Univariate tests

Figure 2 illustrates an uncommon situation, where the desired result is not a maximal power of the test, but a minimal one. The simulated situation is the one of near isotropic tensors, with a progressive change of orientation. As the tensors are all nearly spherical, the uncertainty in their orientation is high, and there should be no noticed difference between them. The tests should not be sensitive to the simulated deformations. However, it can be seen that this

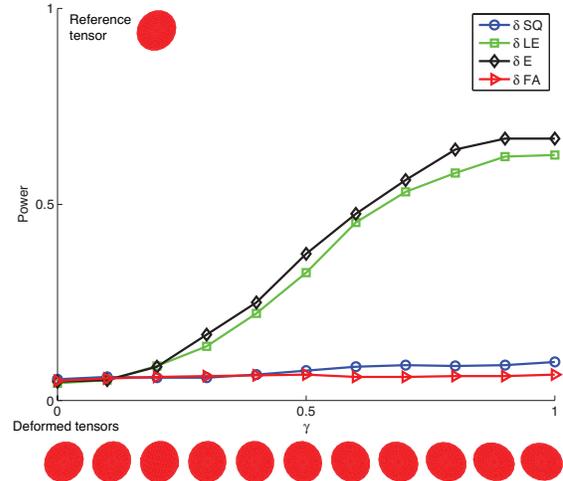


Fig. 2. Power of the statistical tests in a situation of very low anisotropy (the tensors are nearly isotropic). The tests use the spectral quaternion measure (δ SQ, blue circles), the Log-Euclidean measure (δ LE, green squares), the Euclidean measure (δ E, black diamonds) and the Fractional Anisotropy of tensors (δ FA, red triangles). A difference in the diffusion orientation (CO) is simulated. In this case, as illustrated on the figure, the tensors are very similar (all almost spheric) and we argue here that no difference should be noted. However, this is not the case for the Log-Euclidean and Euclidean tests.

is not the case using Euclidean or Log-Euclidean measures. In this case, the spectral-quaternion measure and the test based on Fractional Anisotropy perform better, as they do not detect any difference between the reference and the deformed tensors. It is interesting to note that these curves of power can also be interpreted as curves of sensibility and robustness of the measures. The power of a test is a measure of its sensitivity to the considered deformation. Conversely, flat curves indicate a robustness of the test to a given deformation.

This analysis of robustness is also applicable to the example of Figure 3, showing a difference in orientation in the case of low anisotropy. In this figure, a clear difference can be observed between the Fractional Anisotropy test and the Euclidean and Log-Euclidean tests. As expected, the Fractional Anisotropy test is never significant for this type of difference, showing one more time the limitations of using a unique scalar to represent multivariate data, as the orientation information is totally lost. To the contrary, the Euclidean and Log-Euclidean tests are very sensitive to this difference and thus exhibit strong performance. The spectral-quaternion measure offers an intermediate situation: it is sensitive to the orientation information, but less than the two Euclidean tests. In fact, the orientation term of this measure is weighted by a parameter k , which depends on the anisotropy of tensors. The role of this parameter is to decrease the importance of the orientation term in case of low anisotropy, since in this case, the orientation information is highly uncertain. This explains why the spectral-quaternion measure is not sensitive to the deformations shown in Figure 2. It is important to understand the impact of this parameter of the results. If it is increased,

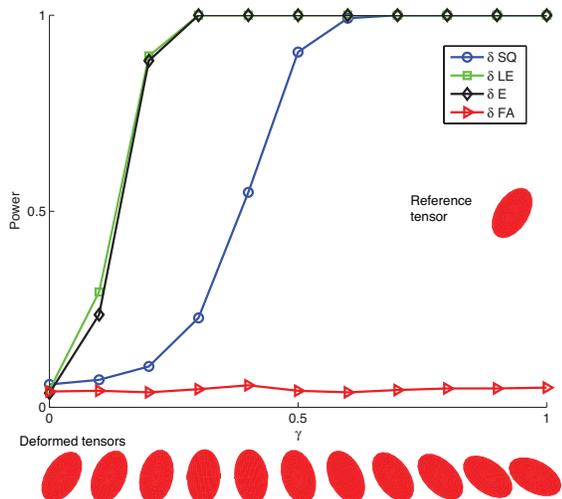


Fig. 3. Power of the statistical tests in a situation of low anisotropy. A difference in the orientation of the tensors is simulated. The tests use the spectral quaternion measure (δ SQ, blue circles), the Log-Euclidean measure (δ LE, green squares), the Euclidean measure (δ E, black diamonds) and the Fractional Anisotropy of tensors (δ FA, red triangles). For $\gamma = 0$, which means no difference between the two references of the groups, the power of the tests is about α . As expected, the test based on the Fractional Anisotropy fails to detect the difference in orientation. The Euclidean and Log-Euclidean tests are very sensitive to this kind of deformation, while the spectral-quaternion measure is between those two situations.

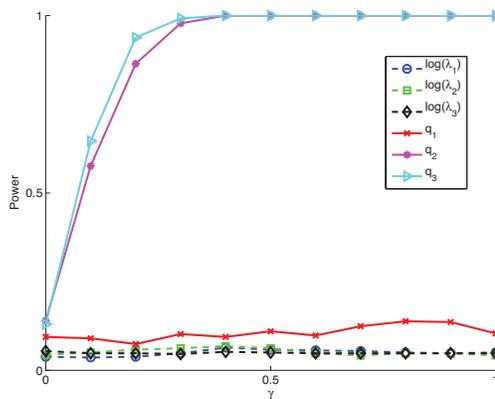
the orientation term will become more important, which will produce an increase of the sensitivity of the measure (*i.e.* an increase of the performance of the test). Depending upon the tradeoff between sensitivity and robustness, the curve of the spectral-quaternion test can be closer to the Fractional Anisotropy one, or to the contrary, closer to the Log-Euclidean test. The fact that this tradeoff can be tuned is of relevance for clinical applications.

B. Multivariate tests

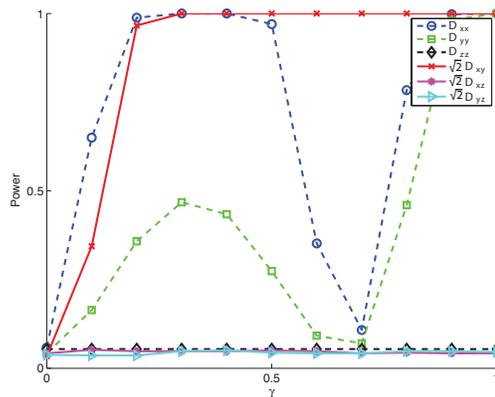
In the following, we will focus on the interpretation of the results of multivariate tests.

Figure 4 illustrates the results of each partial tests for multivariate parametrizations of the tensors, for a simulated change of orientation in a case of high anisotropy. The comparison between geometric parametrization (*top*) and algebraic one (*bottom*) is straightforward. As the geometric parametrization clearly shows that the orientation only has been changed, this interpretation can not be drawn from the results of the Euclidean tests. This information could however be of great importance in clinical studies. In a similar way, the decrease of longitudinal diffusion simulated in Figure 5 is clearly seen with the geometric parametrization, while this is not the case for the Euclidean one. Indeed, for a geometric parametrization, the partial test of the first eigenvalue is the only one to detect a difference. It should be noted that, from $\gamma = 0.9$, the first eigenvalue is very close to the second one, which increases the uncertainty in orientation (this explains why other partial tests become significant). The easy interpretation of the statistical tests using the geometric

parametrization is a desirable feature, which opens the way to many applications. It should be noted that if a unique significance level is needed, the partial tests can be combined using an appropriate function.



(a) Geometric parametrization



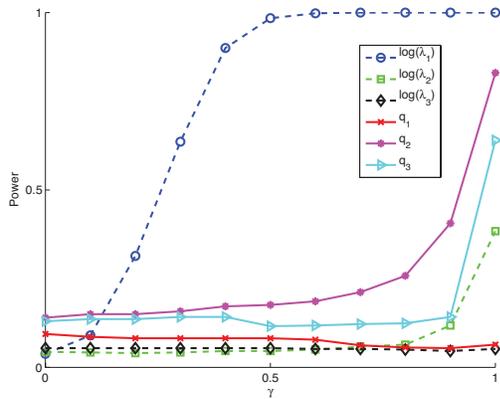
(b) Euclidean parametrization

Fig. 4. Power of each of the partial statistical tests for multivariate parametrizations of the tensors. (a) A geometric parametrization is used. (b) An Euclidean parametrization is used. The simulated situation is a deformation of orientation in a case of high anisotropy. The results of the geometric tests are easily interpretable, as only the partial tests associated to the orientation are performant. To the contrary, the Euclidean tests are poorly understandable.

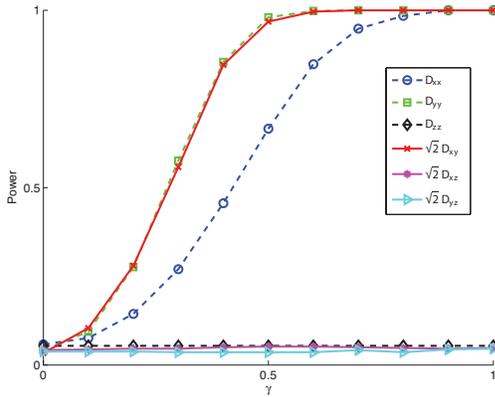
V. CONCLUSIONS

In this work, we have shown that existing methods in the field of group comparison could be advantageously used for the statistical analyses of data lying on Riemannian manifolds. These methods have several advantages as they do not use any assumptions about the distribution of the data (which is seldom known). Moreover, it has been shown that the only specific tool which is needed for this group comparison is an appropriate similarity measure between the data (or a parametrization of them). We have illustrated the computational advantage of basing the permutation test on dispersion rather than means.

Using Diffusion Tensor Images as an example, we have shown how different measures or different parametrizations



(a) Geometric parametrization



(b) Euclidean parametrization

Fig. 5. Power of each of the partial statistical tests for multivariate parametrizations of the tensors. (a) A geometric parametrization is used. (b) An Euclidean parametrization is used. The simulated situation is a decrease of the longitudinal diffusion in a case of high anisotropy. The results of the geometric tests are easily interpretable, as only the partial test associated to the first eigenvalue is performant. To the contrary, the Euclidean tests are difficult to interpret.

of the data can affect the results of the tests. Moreover, two interesting features of the spectral-quaternion measure (and the geometric parametrization associated to this measure) have been highlighted, which both could be relevant for clinical applications.

Both for computational and conceptual reasons, dispersion based permutation tests offer an appealing framework for group comparison of manifold-valued data.

REFERENCES

- [1] J. R. Jensen *et al.*, *Introductory digital image processing: a remote sensing perspective*. Prentice-Hall Inc., 1996, no. Ed. 2.
- [2] P. Basser, J. Mattiello, and D. L. Bihan, “MR diffusion tensor spectroscopy and imaging,” *Biophysical Journal*, vol. 66, no. 1, pp. 259 – 267, 1994.
- [3] M. H. Chappell, J. A. Brown, J. C. Dalrymple-Alford, A. M. Uluğ, and R. Watts, “Multivariate analysis of diffusion tensor imaging data improves the detection of microstructural damage in young professional boxers,” *Magn Reson Imaging*, vol. 26, no. 10, pp. 1398–405, 2008.
- [4] S. Chung, D. Pelletier, M. Sdika, Y. Lu, J. I. Berman, and R. G. Henry, “Whole brain voxel-wise analysis of single-subject serial DTI

- by permutation testing,” *Neuroimage*, vol. 39, no. 4, pp. 1693–705, 2008.
- [5] D. K. Jones, M. R. Symms, M. Cercignani, and R. J. Howard, “The effect of filter size on VBM analyses of DT-MRI data,” *Neuroimage*, vol. 26, no. 2, pp. 546–54, 2005.
- [6] C. B. Goodlett, P. T. Fletcher, J. H. Gilmore, and G. Gerig, “Group analysis of DTI fiber tract statistics with application to neurodevelopment,” *Neuroimage*, vol. 45, no. 1 Suppl, pp. S133–42, 2009.
- [7] A. D. Lee, N. Leporé, F. Lepore, F. Alary, P. Voss, Y. Chou, C. Brun, M. Barysheva, A. W. Toga, and P. M. Thompson, “Brain differences visualized in the blind using tensor manifold statistics and diffusion tensor imaging,” in *Frontiers in the Convergence of Bioscience and Information Technologies, 2007. FBIT 2007*. IEEE, 2007, pp. 470–476.
- [8] A.-L. Fouque, P. Fillard, A. Bargiacchi, A. Cachia, M. Zilbovicius, B. Thyreau, E. Le Floch, P. Ciuciu, and E. Duchesnay, “Voxelwise multivariate statistics and brain-wide machine learning using the full diffusion tensor,” *Med Image Comput Comput Assist Interv*, vol. 14, no. Pt 2, pp. 9–16, 2011.
- [9] A. Grigis, V. Noblet, F. Heitz, F. Blanc, J. de Sèze, S. Kremer, L. Rumbach, and J.-P. Armpach, “Longitudinal change detection in diffusion MRI using multivariate statistical testing on tensors,” *Neuroimage*, vol. 60, no. 4, pp. 2206–21, 2012.
- [10] X. Pennec, “Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements,” *Journal of Mathematical Imaging and Vision*, vol. 25, no. 1, pp. 127–154, 2006.
- [11] P. T. Fletcher and S. Joshi, “Riemannian Geometry for the Statistical Analysis of Diffusion Tensor Data,” *Signal Processing*, vol. 87, pp. 250–262, 2007.
- [12] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, “Statistics on the manifold of multivariate normal distributions: theory and application to diffusion tensor MRI processing,” *Journal of Mathematical Imaging and Vision*, vol. 25, pp. 423–444, 2006.
- [13] I. L. Dryden, A. Koloydenko, and D. Zhou, “Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging,” *The Annals of Applied Statistics*, vol. 3, no. 3, pp. 1102–1123, 2009.
- [14] A. Schwartzman, R. F. Dougherty, and J. E. Taylor, “Group comparison of eigenvalues and eigenvectors of diffusion tensors,” *Journal of the American Statistical Association*, vol. 105, no. 490, pp. 588–599, 2010.
- [15] B. Whitcher, J. J. Wisco, N. Hadjikhani, and D. S. Tuch, “Statistical group comparison of diffusion tensors via multivariate hypothesis testing,” *Magn Reson Med*, vol. 57, no. 6, pp. 1065–74, 2007.
- [16] K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2007.
- [17] T. Nichols and A. Holmes, “Nonparametric permutation tests for functional neuroimaging: a primer with examples,” *Human brain mapping*, vol. 15, no. 1, pp. 1–25, 2002.
- [18] P.-A. Absil, R. Mahony, and R. Sepulchre, “Riemannian geometry of Grassmann manifolds with a view on algorithmic computation,” *Acta Applicandae Mathematicae*, vol. 80, no. 2, pp. 199–220, 2004.
- [19] X. Pennec, P. Fillard, and N. Ayache, “A Riemannian framework for tensor computing,” *International Journal of Computer Vision*, vol. 66, no. 1, pp. 41–66, 2006.
- [20] S. Bonnabel and R. Sepulchre, “Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank,” *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1055–1070, 2009.
- [21] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, “Log-Euclidean metrics for fast and simple calculus on diffusion tensors,” *Magnetic Resonance in Medicine*, vol. 56, pp. 411–421, 2006.
- [22] A. Collard, S. Bonnabel, C. Phillips, and R. Sepulchre, “An anisotropy preserving metric for DTI processing,” *submitted to International Journal of Computer Vision*, 2012.
- [23] P. W. Mielke and K. J. Berry, *Permutation methods: a distance function approach*. Springer Verlag, 2001.
- [24] F. Pesarin, *Multivariate Permutation Tests : With Applications in Biostatistics*. Wiley, 2001.
- [25] H. Boisgontier, V. Noblet, F. Heitz, L. Rumbach, and J.-P. Armpach, “Generalized likelihood ratio tests for change detection in diffusion tensor images: Application to multiple sclerosis,” *Medical Image Analysis*, vol. 16, no. 1, pp. 325–338, 2012.